A new estimator for quantile-oriented sensitivity indices

Thomas Browne

Supervisors : J-C. Fort (Paris 5) & T. Klein (IMT-Toulouse)

Advisors : B. Iooss & L. Le Gratiet (EDF R&D-MRI Chatou)

EDF R&D-MRI Chatou - Université Paris 5

April 19th, 2016, Les Houches

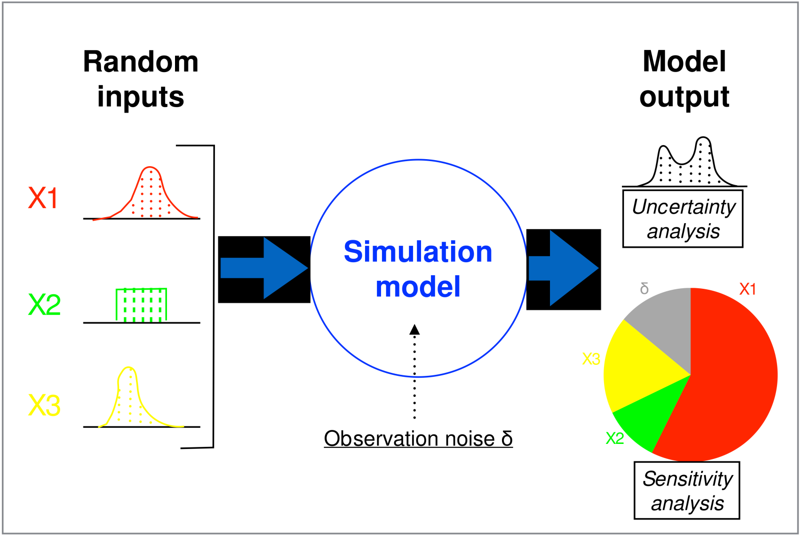
Sensitivity Analysis - Introduction

- Numerical code g.
- Random inputs $(X_1, \ldots, X_d) \sim (f_1, \ldots, f_d) \quad iid$.
- Random output $Y \in \mathbb{R}$ such that

$$Y = g(X_1, \ldots, X_d) .$$

- Main goal : for $i \in \{1, \ldots, d\}$, how does Xi's uncertainty propagate through g ?

Sensitivity analysis - Schema



Sensitivity analysis

- Several potential uses :

  $\longrightarrow$ Better understanding of the model,

  $\longrightarrow$ Neglect X_i's distribution if not influential,

  $\longrightarrow$ Feedback on the inputs - reducing X_i's distribution if too much influential.

- Global analysis : most relevant way ?

Goal-oriented sensitivity analysis

- In practice, Y's distribution does not need to be fully known.
- Choice of a probability feature $\theta(Y)$ (mean, quantiles etc ...) which may be relevant.
- Goal-oriented sensitivity analysis (*GOSA*) [N. Rachdi, 2011] :

  $\longrightarrow$ For $i \in \{1, \ldots, d\}$, quantification of X_i's influence over $\theta(Y)$.

- One by one strategy : condition the code g by X_i and compute $\theta(Y \mid X_i)$

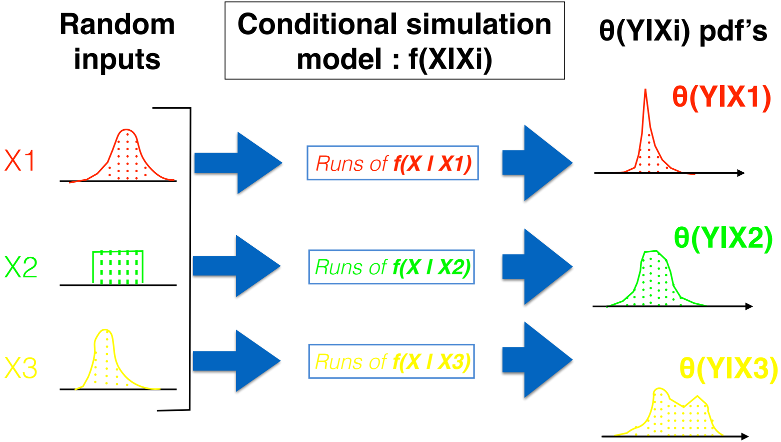
  Set x_i realization of $X_i \longrightarrow g(X_1, \dots, x_i, \dots, X_d) \longrightarrow \theta(Y \mid X_i = x_i)$,

  Condition g by all the possible values x_i, regarding f_i :

  $\longrightarrow \theta(Y \mid X_i)$'s distribution, random variable function of X_i.

GOSA - Schema

Respective influences of each input over θ(Y)



Contrast functions

- Use contrast functions to quantify $\theta(Y \mid X_i)$'s variability.
- Simple contrasts : $\forall (y, \theta) \in \mathbb{R}^2 \quad \varphi(y, \theta) \geq 0$

  quantify a "distance" between two real components.
- Mean contrasts : for Y r.r.v. $\phi_Y(\theta) = \mathbb{E}_Y[\varphi(Y, \theta)]$.
- Y's feature : $\theta(Y) := \arg\min_{\theta \in \mathbb{R}} \phi_Y(\theta)$.

Contrast functions : mean and quantiles

- If $\varphi(y,\theta) = m(y,\theta) =\mid y-\theta \mid^2$:

  therefore $\phi_Y(\theta) = \mathbb{E}_Y[\mid Y-\theta \mid^2]$,

  $\longrightarrow \theta(Y) = \mathbb{E}[Y]$.

- If, for $\alpha \in]0;1[$, $\varphi(y,\theta) = c_\alpha(y,\theta) = (y-\theta)(\alpha - 1_{y\leq\theta})$:

  therefore $\phi_Y(\theta) = \mathbb{E}_Y[(Y-\theta)(\alpha - 1_{Y\leq\theta})]$,

  $\longrightarrow \theta(Y) = q^\alpha(Y)$, α-quantile de Y.

- We focus on $\varphi = c_\alpha$: $\theta(Y) = q^\alpha(Y)$.

  N.B. :

$$\min_\theta \phi\left(\theta \mid X_i = x_i\right) = \mathbb{E}\left[c_\alpha\left(Y, q^\alpha(Y \mid X_i)\right) \mid X_i = x_i\right].$$

Sensitivity analysis with respect to a contrast

- Need to quantify the variability of $\theta(Y \mid X_i)$.
- Sensitivity indices based on contrasts [Fort et al., 2013]

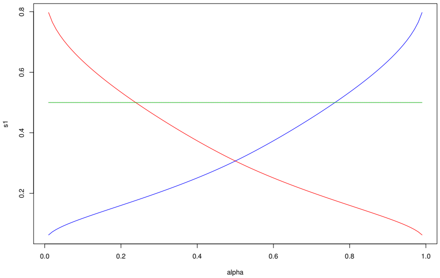
$$S_{\varphi}^{Xi}(Y) = \min_{\theta \in \mathbb{R}} \phi_Y(\theta) - \mathbb{E}\left[\min_{\theta \in \mathbb{R}} \phi_Y(\theta \mid X_i)\right].$$

$\longrightarrow$ quantifies the influence of the input X_i on $\theta(Y)$.

- $S_{\varphi}^{Xi}(Y) \geq 0$.
- We divide $S_{\varphi}^{Xi}(Y)$ by $\min_{\theta \in \mathbb{R}} \phi_Y(\theta)$ so that $0 \leq S_{\varphi}^{Xi}(Y) \leq 1$.

$$S_{c_\alpha}^{i}(Y) = 0 \Leftrightarrow \theta(Y \mid X_i) = \theta(Y) \ a.s.$$
$$S_{c_\alpha}^{i}(Y) = 1 \Leftrightarrow (Y \mid X_i = x_i) = constant(x_i) \ a.s.$$

Sensitivity analysis with respect to a contrast



- $Y = X_1 + X_2$ with $X_1 \sim Exp(1)$ and $X_2 \sim -Exp(1)$ independent.
- $S_m^{X_1} = S_m^{X_2} = 0.5$ (Sobol indices).
- Both inputs are influential on the mean $\mathbb{E}[Y]$!
- $S_{c_\alpha}^{X_1}$: X_1's influence on Y's α-quantile.
- $S_{c_\alpha}^{X_2}$: X_2's influence on Y's α-quantile.
- Sensitivity changes regarding the level of quantile α.

Estimation of the quantile-oriented index

- Goal : from a n-sample $(X_i^1, Y^1), \ldots, (X_i^n, Y^n)$, estimation of

$$S_{c_\alpha}^{Xi}(Y) = \min_\theta \phi_Y(\theta) - \mathbb{E}\left[\min_{\theta\in\mathbb{R}} \phi_Y(\theta \mid X_i)\right].$$

$$= \mathbb{E}\left[c_\alpha(Y, q^\alpha(Y))\right] - \mathbb{E}_{X_i}\left[\min_{\theta\in\mathbb{R}} \mathbb{E}\left[c_\alpha(Y,\theta) \mid X_i\right]\right].$$

- 1st term estimation :

$$\min_\theta \frac{1}{n}\sum_{j=1}^{n} c_\alpha(Y^j, \theta) = \frac{1}{n}\sum_{j=1}^{n} c_\alpha\left(Y^j, \hat{q}^\alpha(Y)\right),$$

where $\hat{q}^\alpha(Y)$ is the classical empirical quantile estimator

$\longrightarrow$ this estimator converges *a.s.*

Estimation for the second term

- Second term : $\mathbb{E}_{X_i}\left[\min_{\theta\in\mathbb{R}} \mathbb{E}\left[c_\alpha(Y,\theta) \mid X_i\right]\right]$.

- Several issues :
  -Double expectation
  -Conditional expectation
  -Minimization problem.

- We use the following asymptotic result [Fan et al., 1994], for x_i any possible realization of X_i :

$$\arg\min_{\theta} \frac{1}{f_i(x_i)} \sum_{j=1}^{n} c_\alpha\left(Y^j,\theta\right) K_{h(n)}\left(X_i^j - x_i\right) \xrightarrow[n\to\infty]{\mathbb{P}} \arg\min_{\theta} \mathbb{E}\left[c_\alpha(Y,\theta) \mid X_i = x_i\right],$$

  where f_i is the *pdf* of X_i with a compact support, K a 2-order positive kernel and $(h(n))_{n\in\mathbb{N}}$ a bandwidth sequence such that $h(n) \to 0$ while $n \times h(n) \to \infty$.

Estimation for the second term

- We define the estimator as :

$$\hat{V}_n = \frac{1}{n} \sum_{k=1}^{n} \frac{1}{k} \min_{\theta} \frac{1}{f_i(X_i^k)} \sum_{j=1}^{k} c_\alpha \left(Y^j, \theta \right) K_{h(k)} \left(X_i^j - X_i^k \right)$$

- Useful points :

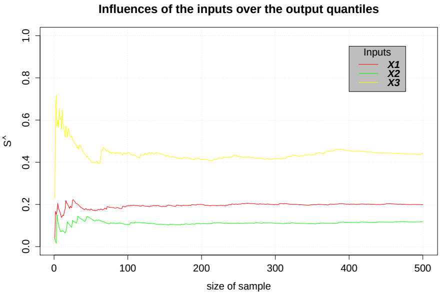
  -As $\left(\theta \mapsto \sum_{j=1}^{k} c_\alpha \left(Y^j, \theta \right) K_{h(k)} \left(X_i^j - X_i^k \right) \right)$ is a piecewise linear function whose "angles" are the $Y^1, ..., Y^k$
  $\longrightarrow$ its minimizer is among $Y^1, ..., Y^k$.

  -$\hat{V}_n$ is built recursively, *ie* if we know it, we also know $V_1, \ldots, V_{n-1}$.

- We prove :

$$\boxed{\hat{V}_n \xrightarrow[n \to \infty]{\mathbb{P}} \mathbb{E}_{X_i} \left[\min_{\theta} \mathbb{E} \left[c_\alpha(Y, \theta) \mid X_i \right] \right]}.$$

Numerical experiments



- Defect detection : wave control through a structure to study.
- Sensitivity analysis over the random defect a_{90}, function of the inputs X, which we detect with a probability of 90%.
- Influence of the inputs over $q^{0.25}(a_{90})$.
- 3 random inputs :
  -X_1 : the thickness of the structure
  -X_2 : the angle of the control
  -X_3 : the depth of the defect.

Conclusions

- Relevant information for the sensitivity analysis - useful alternative to Sobol indices !
- Estimator not so expensive to compute regarding classical estimators in *SA*.
- Convergence criterion for $\hat{V}_n$?
- Perspective : extension to *SA* over random cumulative distribution functions (ouch !)

Sketch of proof for the consistency

- We define a parallel "estimator", V_n, by substituting the minimum, for each $k \in \{1, ..., n\}$, by :

$$\frac{1}{f_i(X_i^k)} \sum_{j=1}^{k} c_\alpha \left(Y^j, q^\alpha \left(Y \mid X^k \right) \right) K_{h(k)} \left(X_i^j - X_i^k \right) .$$

- As we express the increment of $(V_n)_{n \in \mathbb{N}}$, we get :

$$\forall n \in \mathbb{N}^* \quad \frac{V_n - V_{n-1}}{1/n} = (V^* - V_{n-1}) + \varepsilon(n), \quad \text{where } 1/n \text{ is the time-step}$$

and $\varepsilon(n)$ is a "small enough" residual.

- Let us define a real function I that interpolates $(V_n)_{n \in \mathbb{N}}$ such that : $\forall n \in \mathbb{N}\ I(\sum_{k=1}^{n} 1/k) = V_n$. Then :

$$\frac{I(\sum_{k=1}^{n} \frac{1}{k}) - I(\sum_{k=1}^{n-1} \frac{1}{k})}{1/n} \underset{+\infty}{\simeq} \left(V^* - I \left(\sum_{k=1}^{n-1} \frac{1}{k} \right) \right) .$$

Sketch of proof for the consistency

- Under the right conditions, the Kushner-Clark theorem [3] states that, with a probability of 1, g behaves asymptotically like a solution of the associated *ODE** : $I' = V^* - I$
  $\Rightarrow \lim_{t \to +\infty} I(t) = V^*$ *a.s.*, since V^* is the limit of every solution of *ODE**.
- This leads to : $V_n \xrightarrow[n \to \infty]{\mathbb{P}} V^*$.
- By using : $\forall n \in \mathbb{N}\ \hat{V}_n \leq V_n$ and proving $\mathbb{E}\left[|V_n - \tilde{V}_n|\right] \underset{n \to \infty}{\to} 0$
  $\Longrightarrow \quad \hat{V}_n \xrightarrow[n \to \infty]{\mathbb{P}} V^*$.

N. Rachdi
Statistical Learning and Computer Experiments
PhD thesis, Université Paul Sabatier, France, 2011.

J. C. Fort, T. Klein, N. Rachdi
New sensitivity analysis subordinated to a contrast
Communication in Statistics : Theory and Methods, *In press*, 2013.

J. Fan, T. Hu and Y. K. Truong
Robust Non-Parametric Function Estimation
Scandinavian Journal of Statistics,Vol. 21, No. 4, pp. 433-446, 1994.

H. J. Kushner, D. S. Clark
Stochastic Approximations for Constrained and Unconstrained Systems
Springer, Berlin, 1978.

Les Houches, c'est bien.